



Nonnegative tensor CP decomposition of hyperspectral data

Miguel Angel Veganzones, Jérémy E Cohen, Rodrigo Cabral Farias, Jocelyn Chanussot, Pierre Comon

► To cite this version:

Miguel Angel Veganzones, Jérémy E Cohen, Rodrigo Cabral Farias, Jocelyn Chanussot, Pierre Comon. Nonnegative tensor CP decomposition of hyperspectral data. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54 (5), pp.2577-2588. 10.1109/TGRS.2015.2503737 . hal-01134470v2

HAL Id: hal-01134470

<https://hal.science/hal-01134470v2>

Submitted on 24 Nov 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives| 4.0 International License

Nonnegative tensor CP decomposition of hyperspectral data

Miguel A. Veganzones, *Member, IEEE*, Jeremy E. Cohen, Rodrigo Cabral Farias, Jocelyn Chanussot, *Fellow, IEEE*, and Pierre Comon, *Fellow, IEEE*.

Abstract—New hyperspectral missions will collect huge amounts of hyperspectral data. Besides, it is possible now to acquire time series and multiangular hyperspectral images. The process and analysis of these big data collections will require common hyperspectral techniques to be adapted or reformulated. The tensor decomposition, *a.k.a.* multiway analysis, is a technique to decompose multiway arrays, that is, hypermatrices with more than two dimensions (ways). Hyperspectral time series and multiangular acquisitions can be represented as a 3-way tensor. Here, we apply Canonical Polyadic tensor decomposition techniques to the blind analysis of hyperspectral big data. In order to do so, we use a novel compression-based nonnegative CP decomposition. We show that the proposed methodology can be interpreted as multilinear blind spectral unmixing, a higher order extension of the widely known spectral unmixing. In the proposed approach, the big hyperspectral tensor is decomposed in three sets of factors which can be interpreted as spectral signatures, their spatial distribution and temporal/angular changes. We provide experimental validation using a study case of the snow coverage of the French Alps during the snow season.

Index Terms—Hyperspectral, nonnegative tensor decomposition, CP decomposition, big data, compression, time series.

I. INTRODUCTION

IMAGING spectroscopy [1] (*a.k.a.* hyperspectral imaging) is concerned with the measurement, analysis, and interpretation of spectra acquired from a given scene or object [2]. The impact of hyperspectral imaging in remote sensing during the last two decades has been huge [2]–[4]. The interest of the remote sensing community in the subject has grown up to be comparable to that of radar technology, with a clear increasing trend in the former and a stabilization or decrease in the latter [5]. The International Spaceborne Imaging Spectroscopy Technical Committee (ISIS TC) of IEEE GRSS has recently informed of an increasing number of terrestrial space-based civilian imaging spectroscopy missions currently in preparation or in the planning phase [6]. In addition to the Earth Observation (EO) missions, remote sensing hyperspectral imaging has been used for planetological studies [7] and deep space exploration [8]. Also, it has proved

useful for industrial and laboratory applications, using small commercial instruments.

Hyperspectral images (HSI) are usually analysed as a non-negative matrix, $\mathbf{X} \in \mathbb{R}_+^{N \times D}$, where N denotes the number of pixels in the image and D denotes the number of spectral bands. For most of the modern hyperspectral sensors, the number of acquired spectral bands is in the order of hundreds, $O(D) \approx 10^2$, *e.g.*, EnMAP¹, PRISMA² or HYPERION³ sensors capture between 220 and 238 spectral bands covering the spectrum in a range of wavelengths between $0.4\mu\text{m}$ and $2.5\mu\text{m}$. However, it is expected that future sensors will collect thousands of bands, *e.g.*, IASI sensor⁴ captures 8461 spectral bands in wavelengths covering the range $3.62 - 15.5\mu\text{m}$. The number of pixels composing a HSI are usually in the order of hundreds of thousands, $O(N) \approx 10^6$. The high spatial and spectral dimensionality of the HSI makes their analysis very computationally costly. Furthermore, new missions and sensor developments are making it possible to collect time series of hyperspectral data, *e.g.* MODIS mission⁵, and multiangular images, *e.g.* CRISM mission⁶. The huge amount of hyperspectral data that will be delivered in the near future encouraged us to consider hyperspectral image analysis from a big data point of view. These hyperspectral big data will pose new challenges to hyperspectral image analysis. For instance, new techniques to extract the low-rank relevant information will be necessary.

Here, we propose to make use of techniques from tensor analysis [9], [10] (*a.k.a.* multiway or multiarray analysis) to face this challenge. Time series or multiangle hyperspectral big data could be understood as nonnegative tensors, $\mathcal{X} \in \mathbb{R}_+^{N \times D \times T}$, where N , D and T denote the dimensionality of the spatial, spectral and time/angle ways, respectively. One of the most successful techniques to decompose tensors in low-rank terms is the Canonical Polyadic decomposition (CP) [11], sometimes coined Candecomp/Parafac [12]. The CP decomposition could be understood as an extension of the linear unmixing of two-way (spatial and spectral) hyperspectral data [13] to the multilinear unmixing of multiway (more than two ways) hyperspectral tensors. Conventional spectral unmixing aims to decompose a hyperspectral image into the spectral signatures of the materials present in the image and their spatial distributions, known respectively as *endmembers* and

M.A. Veganzones, J.E. Cohen, R. Cabral Farias and P. Comon are with GIPSA-Lab, CNRS, 11 rue des Mathématiques, Grenoble Campus, BP.46, F-38402 St Martin d'Hères cedex, France. J. Chanussot is with GIPSA-lab, Grenoble-INP, F-38402 Saint Martin d'Hères Cedex, France; and, with the Faculty of Electrical and Computer Engineering, University of Iceland, Reykjavik, Iceland. e-mail: {miguel-angel.veganzones, jeremy.cohen, rodrigo.cabral-farias, jocelyn.chanussot, pierre.comon}@gipsa-lab.fr

This work has received funding from the European Research Council under the European Community Seventh Framework Programme FP7/2007-2013 Grant Agreement no. 320594, “DECODA”.

¹<http://www.enmap.org/>

²<http://www.asi.it/en/activity/observation-earth/prisma>

³<http://eo1.usgs.gov/sensors/hyperion>

⁴<http://wdc.dlr.de/sensors/iasi/>

⁵<http://modis.gsfc.nasa.gov/>

⁶<http://crism.jhuapl.edu/>

fractional abundances. Hence, we introduce in Section II the CP decomposition as a technique for blind spectral unmixing of hyperspectral big data. Rank-one factors are expected to be related to spatial abundances, spectral signatures and changes in time/angle. The proposed methodology is blind in the sense that no *a priori* information is needed, for example, when the spectral signatures of the materials in the image are unknown.

Often, conventional spectral unmixing techniques make use of a set of spectral signatures of materials taken on the field or in lab, and collected in a spectral library [14]. Otherwise, the endmembers are estimated from the data using geometrical or statistical spectral unmixing algorithms [13]. Then, the estimation of the spatial fractional abundances is modelled as an optimization problem usually solved by a constrained least squares algorithm. Some recent methods have proposed to jointly estimate the endmembers and their fractional abundances [15], [16]. The proposed methodology allows to jointly unmix big data hyperspectral time series or multiangle acquisitions, decomposing the data into a set of spectral, spatial and time factors, that eventually play the role of endmembers, fractional abundances and time/angle changes.

An important aspect of HSIs is that they relate to physical quantities of the objects composing the scene, *e.g.*, the radiance or the reflectance, which are real nonnegative. In addition, the modelling assumed involves spectral, abundance, or scaling quantities, which are also nonnegative. Hence, it is desirable to impose all terms in the CP decomposition described in Section II to be nonnegative. Moreover, there are some practical issues that should be addressed when computing the nonnegative CP decomposition of hyperspectral tensors. In fact, the big data nature of hyperspectral tensors makes the computational cost of nonnegative CP decomposition algorithms prohibitive for real applications. Thus, it is necessary to develop special-purpose algorithms able to reduce memory requirements and to speed up computations, while enforcing nonnegativity of the CP decomposition. Recently, a solution to this issue has been proposed in [17], as well as two compression-based nonnegative CP decomposition algorithms. This issue is addressed in the paper.

A. Related work

The use of tensor analysis appeared in the hyperspectral community only recently. Some recent works [18]–[22] made use of these techniques to compress, denoise and extract spectral-spatial features for classification. In general, the use of the Tucker [10] decomposition, where the factors are constrained to be orthogonal, has been favoured with respect to the CP tensor decomposition, since its application is more related to the above topics (*i.e.*, denoising and compression). When the CP decomposition is used, we have noted that authors propose to use both spatial support dimensions (lines and columns) as two different ways of the tensor. Contrary to that, here we propose to consider the spatial dimensions as a single way, by vectorizing the image. The reason is that, if the two spatial support dimensions are considered as two tensor ways, then, the tensor is not low rank any more. We argue that the capability of explaining the data using only a few

physically interpretable factors is of paramount importance for the present work and spectral unmixing topic in general and then, considering the spatial information as a single way is required.

Very recently, we sketched the proposed approach in [23] and provided some experimental validation using synthetic data. Here, we extend this preliminary work by giving an extensive theoretical support and additional experimental validation using real data.

B. Contribution

We propose the use of compression-based nonnegative CP decomposition algorithms to analyze big hyperspectral data tensors, *i.e.* hyperspectral time series or hyperspectral multi-angle acquisitions. The proposed methodology could be interpreted as an extension of the linear spectral unmixing to higher orders, that is, a multilinear spectral unmixing. Although the proposed nonnegative CP decomposition is not a model derived from physical interpretations, we give experimental evidence that the tensor decomposition provides a physically interpretable multilinear spectral unmixing model, and that the estimated factors are correlated to physical quantities such as the spectral signatures of the materials and their fractional abundances. The provided quantitative and qualitative evidence of the validity of the proposed methodology was obtained using a dataset collected by the MODIS hyperspectral sensor during the 2012 snow season in the French Alps.

The remainder of the paper is organized as follows. In Sec. II we overview the topic of tensor decompositions, *a.k.a.* multiway analysis. In Sec. III we propose an adaptation of the CP decomposition to hyperspectral big data. In Sec. IV we provide experimental evaluation of the proposed methodology using a case study of a hyperspectral time series of MODIS acquisitions. Finally, we give some conclusions in Sec. V.

II. TENSOR DECOMPOSITION / MULTIWAY ANALYSIS

A. Rank revealing decomposition

For our need, a third order tensor \mathcal{X} of size $N \times D \times T$ will be merely assimilated to its representation by a three-way array \mathcal{X}_{ijk} , $1 \leq i \leq N$, $1 \leq j \leq D$, $1 \leq k \leq T$. Note that tensors of order two are then just matrices. Among the set of tensors, the class of so-called *decomposable* tensors play a central role. These tensors, also sometimes called *pure* or *elementary*, may be seen as a discretization of functions whose variable separate:

$$\mathcal{D}_{ijk} = a_i b_j c_k.$$

It is worth noting that there are obviously infinitely many ways to write a decomposable tensor as a product of single-index components, because of scaling indeterminacies. In fact, one can replace $a_i b_j c_k$ by $(a_i/\alpha)(b_j/\beta)(\alpha\beta c_k)$ for any pair of nonzero scalars (α, β) . Any finite tensor can be written as a finite sum of decomposable tensors:

$$\mathcal{X} = \sum_{r=1}^R \lambda_r \mathcal{D}(r), \quad (1)$$

where λ_r is a strictly positive real. The *rank* of tensor \mathcal{X} is defined as the minimal number R of terms necessary for the equality above to hold exactly. In that case, expression (1) is called the *Canonical Polyadic* (CP) decomposition of \mathcal{X} . Decomposable tensors have then, by definition, a rank equal to 1. This definition of rank is consistent with the usual definition of matrix rank, when tensors of order two are considered. In practice, it may be useful to impose tensors $\mathcal{D}(r)$ to be of unit norm. With this normalization, λ_r can be compared to singular values in the case of tensors of order two.

A first consequence of the nonnegativity constraint is that it has an impact on the value of the tensor's rank. In fact, the (matrix or tensor) rank computed in \mathbb{R}^+ may be strictly larger than the rank computed in \mathbb{R} [24], [25]. For this reason, when *all quantities* involved in the CP decomposition (namely a_i , b_j and c_k) are real nonnegative, the minimal value of R is called the *nonnegative rank*.

B. Factor matrices

Another writing of the CP decomposition makes it more explicit that the data tensor \mathcal{X} is related to a diagonal tensor \mathcal{L} via a multilinear transform. To see this, write $\mathcal{D}_{ijk}(r) = A_{ir}B_{jr}C_{kr}$, where factor matrices \mathbf{A} , \mathbf{B} and \mathbf{C} are of size $N \times R$, $D \times R$ and $T \times R$, respectively. Then, it is easy to show that (1) can be rewritten as:

$$\mathcal{X}_{ijk} = \sum_{r=1}^R A_{ir}B_{jr}C_{kr}\lambda_r. \quad (2)$$

Now define \mathcal{L} as the diagonal tensor of size $R \times R \times R$ containing the R scaling factors λ_r . Then, we agree to denote compactly the polyadic decomposition (2) as

$$\mathcal{X} = (\mathbf{A}, \mathbf{B}, \mathbf{C}) \cdot \mathcal{L}. \quad (3)$$

The major difference between (1) and (2-3) is that the latter made use of one writing of $\mathcal{D}(r)$ among others, and hence cannot pretend to be unique because of this arbitrary choice. On the other hand, this writing is unavoidable when implementing numerical algorithms, hence its importance. This issue will be addressed in Subsection II-D.

C. Approximation

In practice, the data tensor is subject to modelling errors or measurement noise, and it is convenient to find its best rank- R approximation by minimizing the following objective function

$$\Upsilon(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathcal{L}) = \|\mathcal{X} - (\mathbf{A}, \mathbf{B}, \mathbf{C}) \cdot \mathcal{L}\|, \quad (4)$$

for some well chosen norm, instead of attempting to compute the exact CP decomposition (2). It is now known that tensors of order 3 or larger do not always admit a rank- R approximate, when $R > 1$, especially in \mathbb{R} or \mathbb{C} . But fortunately, it has been shown in [25] that this obstacle does no longer holds for nonnegative tensors, and that the problem is well-posed in \mathbb{R}^+ : best lower nonnegative rank approximates always exist.

D. Uniqueness

The exact CP decomposition (1) is unique if the rank R is not too large. In fact, several sufficient upper bounds have been proposed in the literature [26], [27]. In particular, any tensor of rank smaller than or equal to the so-called Kruskal's bound R_o below has a unique CP decomposition:

$$R_o = \left\lfloor \frac{N + D + T - 2}{2} \right\rfloor. \quad (5)$$

However, uniqueness results available in the literature apply only to the exact CP decomposition, and not to an approximation like (4), which must unavoidably be computed in practice. Therefore, we shall also rely on a recent result of [28], which states that almost every nonnegative tensor of nonnegative rank larger than R admits a unique best approximate of nonnegative rank R , which in turn has a unique CP decomposition if condition $R \leq R_o$ holds true.

In practice, the number R of nonnegative rank-one factors is always much smaller than the bound R_o given in equation (5), which ensures almost surely uniqueness of the best rank- R approximate. For instance, in the experiment described in subsection IV-C2, we had $R = 8$ rank-one factors, whereas the bound is $R_o = 2424$.

III. HYPERSPECTRAL BIG DATA CP DECOMPOSITION

A. Extending the linear unmixing model

The linear unmixing model is the simplest and most widely used model for recovering spectra and abundances of a scenery from a hyperspectral image. By imposing some constraints on the factors, the bilinear decomposition becomes unique, and the recovered factors bear physical meaning as actual spectra and abundances. In fact, the linear unmixing model is a particular case of CP decomposition where the data tensor \mathcal{X} is a matrix, i.e. $T = 1$, and additional constraints are added. Thus in the presence of a third variability source, say time or angle, it is natural to study the constrained CP decomposition of multiway hyperspectral data.

However, unlike its bilinear counterpart, the multiway model cannot be *a priori* deduced from a physical reasoning. Indeed, a multiplicative model will not explain the sophisticated temporal or angular evolution of different materials which includes appearance/disappearance and strong nonlinearities. For this reason, applying the CP decomposition to \mathcal{X} should rather be understood as decomposing the data on rank-one meaningful linear subspaces maximizing the explained variance, very much like a non-orthogonal PCA. Yet, knowing the bilinear model bears physical interpretability. We hope the recovered factors \mathbf{A} , \mathbf{B} and \mathbf{C} can be interpreted respectively as spatial, spectral and time/angle ways of R materials. Moreover, the same material might be decomposed in two or more rank-one factors to capture its complex variability. Also, nonnegativity constraints should be imposed on all three factor matrices to ensure both algorithmic convergence and physical interpretability.

B. Nonnegative CP approximation

It is well known in the optimization community that computing the nonnegative CP decomposition of a positive tensor is a difficult problem. Given a multiway data set \mathcal{X} , we want to solve the following minimization problem:

$$\begin{aligned} \operatorname{argmin} \quad & \|\mathcal{X} - (\mathbf{A}, \mathbf{B}, \mathbf{C}) \cdot \mathcal{L}\|_F^2 \\ \text{w.r.t. } & \mathbf{A}, \mathbf{B}, \mathbf{C} \\ \text{s.t. } & \mathbf{A} \succeq 0, \mathbf{B} \succeq 0, \mathbf{C} \succeq 0, \end{aligned} \quad (6)$$

where $\|\cdot\|_F$ denotes the Frobenius norm. This problem is highly non-convex, yet many algorithms provide rather precise but costly computation, and these algorithms can be divided into two main classes:

- *All-at-once gradient-based descent*, e.g. [29]: all CP parameters are updated at the same time using a gradient scheme (standard or conjugate gradient) and nonnegativity constraints are implemented through barriers or soft penalizations.
- *Alternating minimization*: the cost function is minimized in an alternating way for each factor (\mathbf{A} , \mathbf{B} or \mathbf{C}) while the others are fixed. The most commonly used method for nonnegative CP decomposition is alternating nonnegative least squares (ANLS), e.g. [9].

C. Handling big data through compression

Now in hyperspectral imaging, as stated in the introduction, the usual dimensions of the data tensor are huge. In this setting, the workhorse techniques described briefly above can fail to handle all the data within the memory of the computer, or can converge very slowly.

Large tensors decomposition is actually a hot topic in the tensor decomposition area, especially when constraints are included in the optimization problem. An approach to handle large tensor decomposition is through the use of compression. The general idea is that the original data array \mathcal{X} can be equivalently represented by one or a few arrays \mathcal{X}_c with reduced dimensions $N_c \times D_c \times T_c$. The compressed tensor is then decomposed by solving

$$\begin{aligned} \operatorname{argmin} \quad & \Upsilon = \|\mathcal{X}_c - (\mathbf{A}_c, \mathbf{B}_c, \mathbf{C}_c)\|_F^2 \\ \text{w.r.t. } & \mathbf{A}_c, \mathbf{B}_c, \mathbf{C}_c, \end{aligned} \quad (7)$$

where \mathbf{A}_c , \mathbf{B}_c , \mathbf{C}_c are compressed versions of the original factor matrices, with reduced number of rows N_c , D_c , T_c , but the same number of columns R . For simplification purposes the diagonal matrix of scalings is absorbed in \mathbf{C}_c . Note that, after the compressed factors are obtained, a decompression operation is carried out to recover the factors in the original dimensions.

In the literature, two main approaches have been proposed to compress the tensor array. The first proposed approach [30, p. 92] is through an approximation of the High Order Singular Value Decomposition (HOSVD). The HOSVD [31] approximates the original data in the following way

$$\mathcal{X}_{ijk} \approx \sum_{lmn}^{N_c, D_c, T_c} U_{il} V_{jm} W_{kn} [\mathcal{X}_c]_{ijk}, \quad (8)$$

or using the same notation as in the CP model

$$\mathcal{X} \approx (\mathbf{U}, \mathbf{V}, \mathbf{W}) \mathcal{X}_c, \quad (9)$$

where \mathbf{U} , \mathbf{V} and \mathbf{W} are matrices with orthogonal unit-norm columns. In practice, these matrices are obtained by truncating the first N_c , D_c and T_c left singular vectors of the 3 unfoldings of \mathcal{X} (unfoldings are different concatenations of matrix slices of the tensor, see e.g [10] and references therein). Note that, in this case, the compressed and uncompressed factors are related in the following way:

$$\mathbf{A} \approx \mathbf{U} \mathbf{A}_c, \quad \mathbf{B} \approx \mathbf{V} \mathbf{B}_c, \quad \mathbf{C} \approx \mathbf{W} \mathbf{C}_c, \quad (10)$$

which shows that \mathbf{U} , \mathbf{V} and \mathbf{W} can be seen as decompression operators.

The second approach, called PARACOMP [32], consists in generating multiple compressed tensors through multiple U^i , V^i and W^i , with $i \in \{1, \dots, I\}$, but in this case, these matrices are neither deterministic, neither data dependent. The vectors forming the columns of these matrices are independently drawn from an independent and identically distributed multivariate Gaussian distribution. Since there are multiple compressed factors \mathbf{A}_c^i , the uncompressed factors are obtained by merging the results. After correcting the permutations between different \mathbf{A}_c^i , the results are merged through the solution of a linear system of equations. For example, factor \mathbf{A} is obtained by solving

$$[U^1 \dots U^I]^\top \mathbf{A} = \begin{bmatrix} \mathbf{A}_c^1 \\ \vdots \\ \mathbf{A}_c^I \end{bmatrix}. \quad (11)$$

D. Algorithms for multiway hyperspectral data

The complexity (per iterate), i.e. number of multiplications, to retrieve the compressed factors using iterative algorithms such as gradient descent or alternating least squares (ALS) is of order $\mathcal{O}(N_c D_c T_c R)$ for the HOSVD approach and $\mathcal{O}(N_c D_c T_c R I)$ for the PARACOMP approach. Therefore, if $N_c \ll N$, $D_c \ll D$ and $T_c \ll T$, both approaches may decrease substantially the complexity required for the unconstrained CP decomposition, thus solving the big data issue, which is fundamental to the extension of the linear mixing model to a multilinear one. Still an important issue persists, *how to deal with nonnegativity constraints?* Both, the hyperspectral data and the model parameters are nonnegative. The approaches presented above cannot directly handle nonnegative constraints, since nonnegativity cannot be imposed directly in the compressed space. Therefore, specialized algorithms are needed to deal with nonnegativity in these two approaches.

In the HOSVD approach, we need to solve (7) under the following constraints

$$\mathbf{U} \mathbf{A}_c \succeq 0, \quad \mathbf{V} \mathbf{B}_c \succeq 0, \quad \mathbf{W} \mathbf{C}_c \succeq 0. \quad (12)$$

To solve this problem, adapted instances of the two main classes of algorithms for nonnegative CP decomposition are proposed in [17]. In the all-at-once gradient descent setting, the

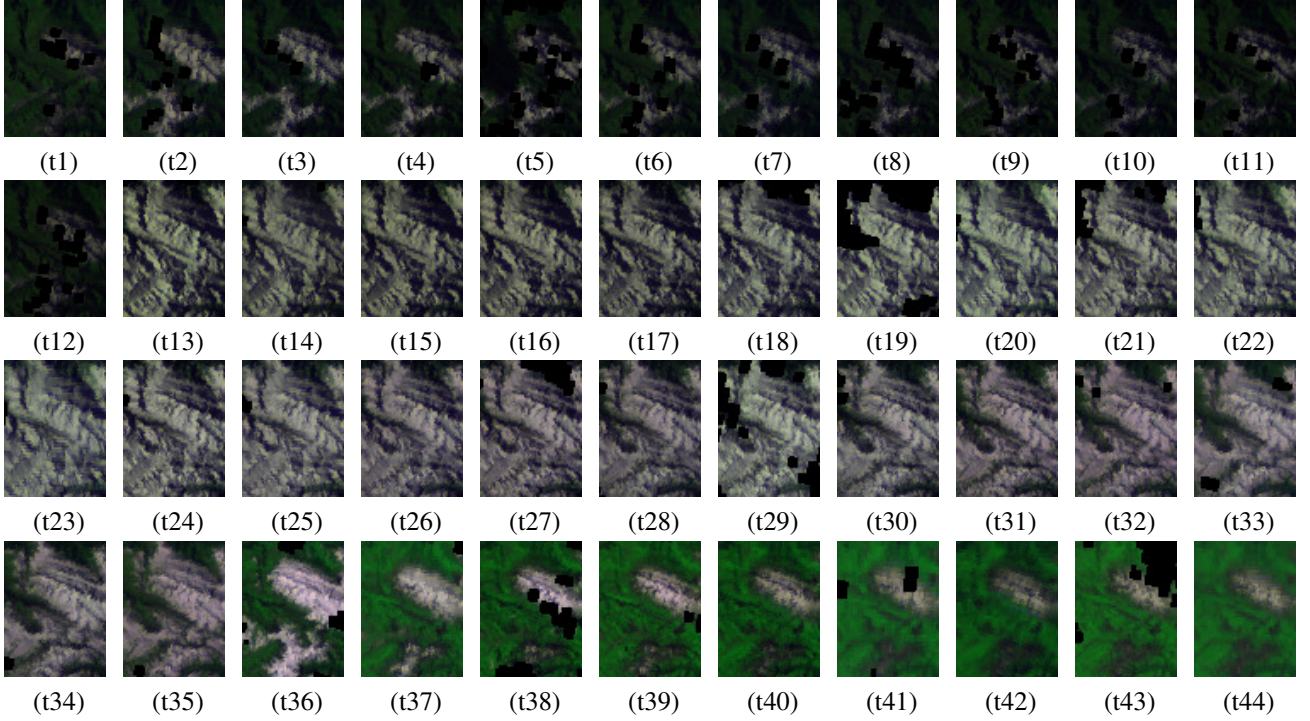


Fig. 1. False color images of the 44 time acquisitions (cloud pixels are depicted in black).

Compressed Conjugate Gradient (CCG) algorithm is proposed, while a modification of ANLS, Projected and Compressed ALS (ProCo ALS), is presented in the alternating setting. A solution similar to ProCo ALS is also proposed in [33].

In the PARACOMP approach, an ADMM projection algorithm is proposed to add nonnegativity constraints [34]. But it has been noted in [33] that this approach is very sensitive to noise, which is a huge drawback for hyperspectral applications. Therefore, in what follows we will briefly explain and apply only the HOSVD-based nonnegative CP decomposition algorithms: CCG and ProCo ALS.

1) *CCG*: to include the nonnegative constraints in the minimization of $\Upsilon(\mathbf{A}_c, \mathbf{B}_c, \mathbf{C}_c)$ soft penalization terms $f([U\mathbf{A}_c]_{ij})$ are added in order to increase the objective function whenever $[U\mathbf{A}_c]_{ij}$ are negative. Stacking all elements of the factor matrices in a single vector $\boldsymbol{\theta} = \text{vec}(\mathbf{A}_c^\top, \mathbf{B}_c^\top, \mathbf{C}_c^\top)$, the modified objective function becomes

$$\Upsilon_p = \Upsilon(\boldsymbol{\theta}) + \frac{1}{R(N+D+T)} \left[\sum_{i,j} f([U\mathbf{A}_c]_{ij}) + \dots \right],$$

where f is a sigmoidal function, for example a hyperbolic tangent function. Then, applying the conjugate gradient method [35, p. 101] to this function gives the CCG method. At iterate $k+1$, the CCG update $\hat{\boldsymbol{\theta}}^{k+1}$ is given by

$$\hat{\boldsymbol{\theta}}^{k+1} = \hat{\boldsymbol{\theta}}^k + \mu_k \mathbf{s}^k, \quad (13)$$

where μ_k is the step size given by backtracking line search [35, p. 37] and \mathbf{s}^k is the conjugate gradient direction, for example

using the Polak-Ribière method [35, p. 122]⁷, this direction is given by

$$\mathbf{s}^{k+1} = \mathbf{p}^{k+1} + \frac{(\mathbf{p}^{k+1})^\top (\mathbf{p}^{k+1} - \mathbf{p}^k)}{\|\mathbf{p}^k\|^2} \mathbf{s}^k, \quad (14)$$

where \mathbf{p}^k denotes the negative gradient of Υ_p at the point $\hat{\boldsymbol{\theta}}^k$.

2) *ProCo ALS*: in the alternating setting, to update each factor we need to solve a least squares problem with linear inequality constraints. The idea of ProCo ALS is to carry out an approximate projection of the unconstrained least squares (LS) solution on the set of feasible solutions. This approximate projection is given in three steps: first the factor is decompressed (D), then the decompressed factor is projected onto the nonnegative orthant (P), and, finally, the result of the projection is recompressed (R). The four steps of each ProCo ALS iteration to update the factor \mathbf{A}_c , denoted here $\hat{\mathbf{A}}_{pc}^{k+1}$, are detailed below:

$$\begin{aligned} \text{LS:} \quad & \hat{\mathbf{A}}_c^{k+1} = \mathbf{X}_c^{(1)} (\hat{\mathbf{C}}_{pc}^k \odot \hat{\mathbf{B}}_{pc}^k)^\dagger \\ \text{Approx. proj.} \quad & \left\{ \begin{array}{l} \text{D: } \hat{\mathbf{A}}^{k+1} = \mathbf{U} \hat{\mathbf{A}}_c^{k+1} \\ \text{P: } [\hat{\mathbf{A}}^{k+1}]_+ := \max(\mathbf{0}, \hat{\mathbf{A}}^{k+1}) \\ \text{R: } \hat{\mathbf{A}}_{pc}^{k+1} = \mathbf{U}^\top [\hat{\mathbf{A}}^{k+1}]_+ \end{array} \right. \end{aligned} \quad (15)$$

where \odot denotes the Khatri-Rao product, \dagger indicates the pseudoinverse, $\mathbf{X}_c^{(1)}$ is the unfolding in the first way of \mathbf{X}_c and $\max(\cdot, \cdot)$ denotes the element wise maximum. To obtain the

⁷Note that in [17], the Fletcher-Reeves implementation was used. In (14), we use the Polak-Ribière implementation instead, since its convergence was observed to be faster in practice.

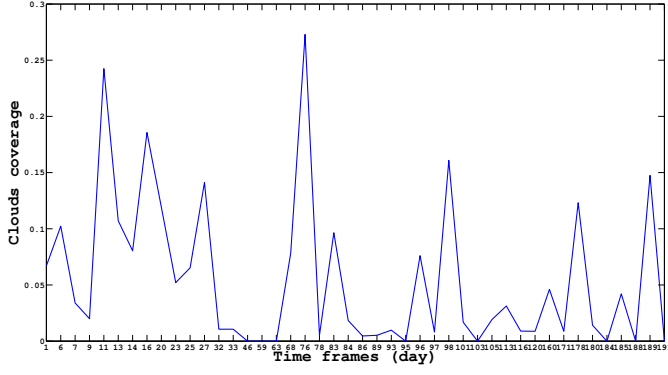


Fig. 2. Cloud coverage of the 44 time acquisitions indicated by the ratio of cloud pixels to the total number of pixels (0.00 indicates no clouds, 1.00 indicates fully covered by clouds).

updates $\hat{\mathbf{B}}_{pc}^{k+1}$ and $\hat{\mathbf{C}}_{pc}^{k+1}$, the same procedure is applied with the appropriate unfolding of \mathcal{X}_c and decompression operators.

Observe that, for both algorithms, the additional complexity (per iterate) which is required to introduce the constraints in the uncompressed space is of order $\mathcal{O}[(NN_c + DD_c + TT_c)R]$ (due to matrix products $\mathbf{U}\hat{\mathbf{A}}_c$, $\mathbf{V}\hat{\mathbf{B}}_c$ and $\mathbf{W}\hat{\mathbf{C}}_c$). Even though these terms may dominate the complexity when $N_c \ll N$, $D_c \ll D$ and $T_c \ll T$, they are still much smaller than $\mathcal{O}(NDTR)$, which is the complexity for the uncompressed approaches.

IV. STUDY CASE: HYPERSPECTRAL TIME SERIES

A. Dataset

The dataset is a subset of a longitudinal daily acquisition of MODIS hyperspectral sensor for the same scene in the Alps (France) during the 2012 snow season. The data has been pre-processed to improve the spatial resolution to 250m. From the original dataset we have selected 44 acquisitions with a cloud presence lower than 30%. Each image is of 80×60 pixels size with seven spectral bands measuring the radiance at the sensor.

Fig. 1 shows a false color image of the 44 acquisitions. It can be appreciated that some permanent snow/ice lies on the top of the mountain chain, and how the snow covers the vegetation in the middle of the season to finally melt and disappear by the end of the season. Fig. 2 depicts the cloud ratio for each of the 44 images. Most of the images are partially cloudy, and the pixels covered by clouds are considered as missing data. Since it is unlikely that the missing data are located in the same pixel location on consecutive acquisitions, we have interpolated the missing values using the pixel values of time adjacent images.

B. Experimental methodology

On one hand, we compared the application of both compression-based CP decomposition algorithms, CCG and ProCo ALS, to the state-of-the-art ANLS CP decomposition algorithm. In order to do that, we arranged the dataset in a tensor of $4800 \times 7 \times 44$ dimensions, corresponding to the spatial, spectral and time ways, respectively. The tensor size

is convenient since it reflects the usual unbalance on the tensor dimensions, where spatial dimension is much larger than the dimensionality of the other tensor ways, while still allowing the computation of the ANLS algorithm for sake of comparison. We have also compared the proposed approach to the conventional use of spectral unmixing, independently applied to each time acquisition. For that sake, we employed on-field spectral measurements of eight materials (see Fig. 3), here on termed as endmembers according to the unmixing literature: *medium snow*, *glacier snow*, *old coarse*, *ice*, *debris*, *rocks*, *rain forest* and *pasture*. To estimate the abundances, we made use of the fully constrained least-squares unmixing (FCLSU), implemented using the sparse unmixing by variable splitting and augmented Lagrangian (SUnSAL) algorithm [36] where we have enforced the non-negativity and abundance sum-to-one constraints.

We run 50 Monte Carlo runs for each of the algorithms for a set of different rank values in the range $R \in [5, 15]$. For the compression-based CP algorithms, we manually set the dimensions of the compressed tensor, \mathcal{X}_c , to $175 \times 7 \times 25$. This way, we severely compressed the spatial and time ways, aiming to show that the proposed compression-based CP decomposition approach could still obtain a small reconstruction error and estimate physically meaningful factors when the compression ratio is very high. The comparison was done in terms of average normalized root mean squared error (nRMSE), between the original tensor, \mathcal{X} , and the tensor reconstructed from the CP factors, $\hat{\mathcal{X}}$

$$\widehat{\text{nRMSE}}(\mathcal{X}, \hat{\mathcal{X}}) = \sqrt{\frac{1}{NDT} \sum_{i=1}^N \sum_{j=1}^D \sum_{k=1}^T \left(\frac{\mathcal{X}_{ijk} - \hat{\mathcal{X}}_{ijk}}{\|\mathcal{X}\|_F} \right)^2}.$$

On the other hand, we assessed the physical interpretability of the estimated CP factors. We compared the estimated spectral factors, $\hat{\mathbf{B}}$, to the aforementioned on-field spectral measurements. We made use of the spectral angular distance, d_{SAD} , to compare the pair-wise angular distances between the spectral factors and the endmembers:

$$d_{\text{SAD}}(\hat{\mathbf{b}}, \mathbf{e}) = \arccos \left(\frac{\hat{\mathbf{b}}^T \mathbf{e}}{\|\hat{\mathbf{b}}\| \|\mathbf{e}\|} \right), \quad (16)$$

where $\hat{\mathbf{b}} \in \mathbb{R}^D$ denotes a spectral CP factor and $\mathbf{e} \in \mathbb{R}^D$ denotes an endmember. In order to assess the interpretability of the spatial factors, we calculated the linear Pearson correlation of the estimated spatial factors, $\hat{\mathbf{A}}$, to the abundances estimated by a conventional FCLSU, individually applied to each of the time images, using the previously mentioned eight endmembers. Finally, we give some interpretation of the time factors, \mathbf{C} , in terms of the snow season evolution.

C. Results

First, we give some overall results taking into account the performance of the competing nonnegative CP algorithms for all the evaluated rank values. The obtained results suggest that $R = 8$ is an appropriate value for this data tensor rank. Thus, we continue showing some detailed results for the best runs among the experiments with rank $R = 8$, that is the

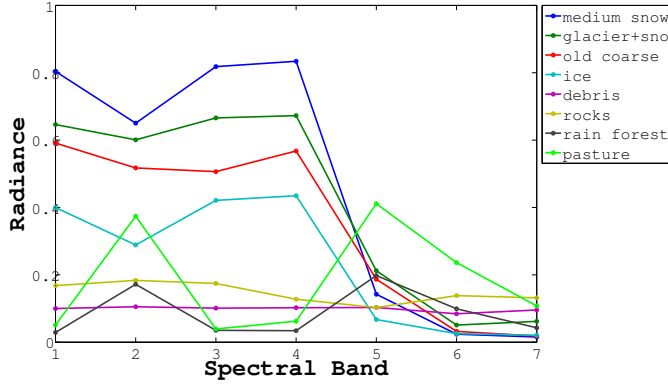


Fig. 3. On-field spectral measurements (endmembers).

runs with minimum average normalized RMSE for each CP decomposition algorithm.

1) *Overall results:* Fig. 4 shows a boxplot of the average normalized RMSE of the 50 Monte Carlo runs for the different CP decomposition algorithms and rank values. The flat black line depicts the average normalized RMSE obtained by the conventional FCLSU spectral unmixing using the eight library endmembers. Overall, the compression-based algorithms achieve a performance similar to the state-of-the-art ANLS. This supports the use of the CCG and ProCo ALS algorithms to perform a CP decomposition of big hyperspectral data since the computational time employed by these two algorithms is several orders of magnitude smaller than the ANLS. The ProCo ALS algorithm shows less variability in the reconstruction error than CCG for small rank values. Results obtained for the rank value $R = 8$ suggests that this is a proper rank value for the CP decomposition. The CCG and ProCo ALS present small error variability for rank values $R \geq 8$ compared to the ANLS algorithm, which seems to be more sensitive to an over-estimation of the tensor rank. Compared to the error obtained by the FCLSU approach, the compressed CP decomposition algorithms achieve similar or slightly better results, whilst the decomposition is more compact, i.e. the FCLSU returns 44 abundance maps for each endmember while the CP decomposition outputs $R < 8 \times 44$ abundance maps and time factors.

Fig. 5 shows the average angular distance between the eight library endmembers and the spectral factors obtained by the 50 Monte Carlo runs of each of the CP decomposition algorithms. Since we want to know if the spectral factors resemble any of the library endmembers, for each run, we select the minimum angular distance from any of the estimated spectral factors to each of the eight library endmembers, and then we compute the mean among the 50 runs. All the three algorithms present the same trends. The four endmembers related to snow/ice (*medium snow*, *glacier snow*, *old coarse*, *ice*) have very small average angular distances for all rank values, meaning that these materials have been detected by the CP decomposition. The *rocks* and *debris* endmembers have high angular values meaning that the CP decompositions do not estimate any spectral factor that resembles these two materials. The *vegetation* and *pasture* are specially present for

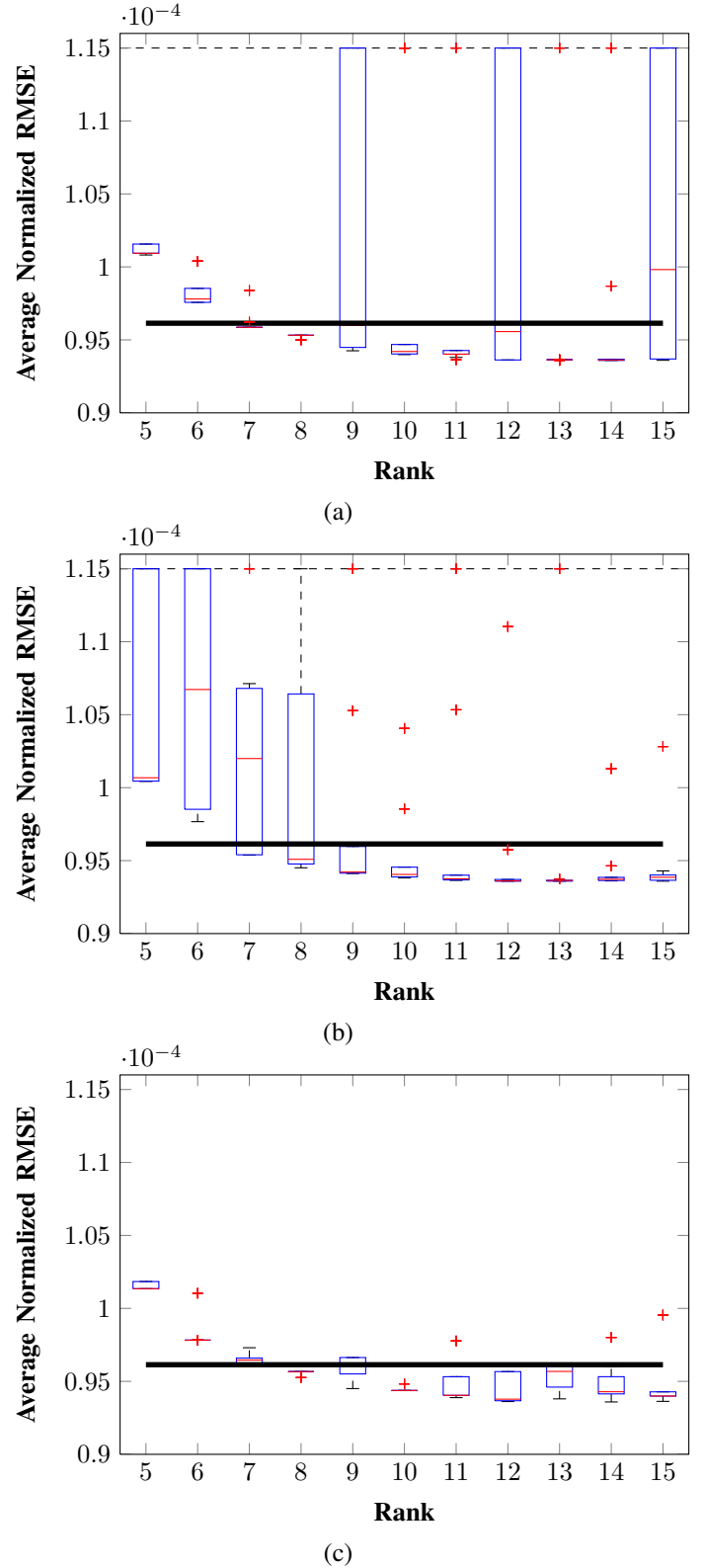
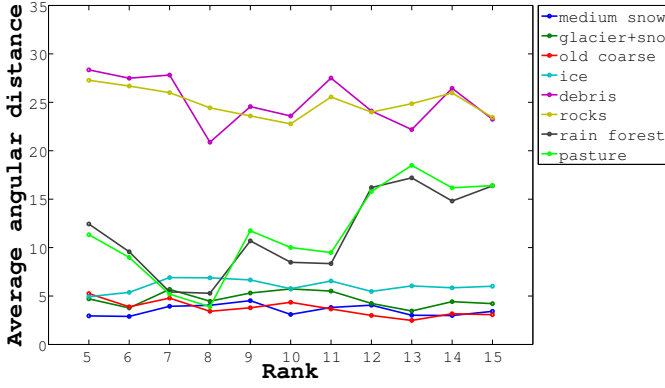
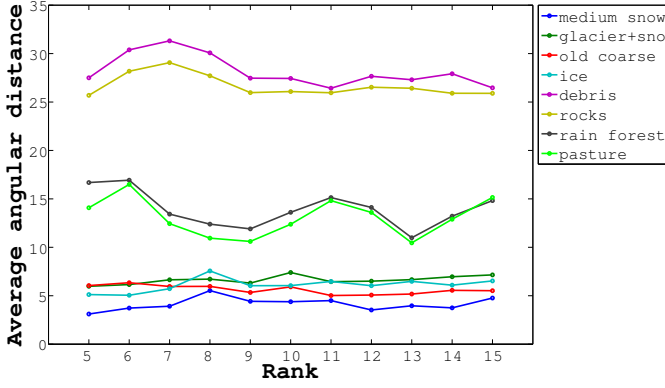


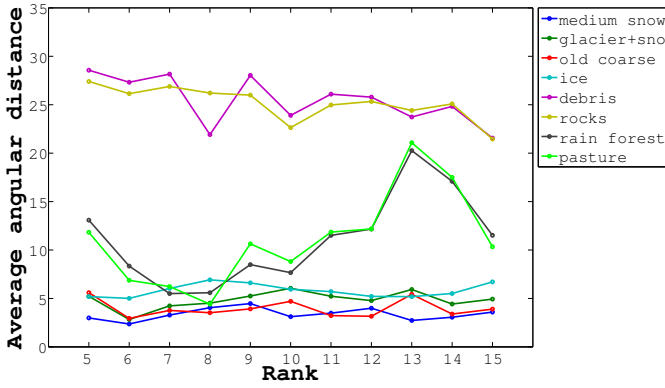
Fig. 4. Average Normalized RMSE obtained by the three competing CP decomposition algorithms: (a) ANLS, (b) CCG and (c) ProCoALS. The black line depicts the error obtained by the conventional FCLSU.



(a)



(b)

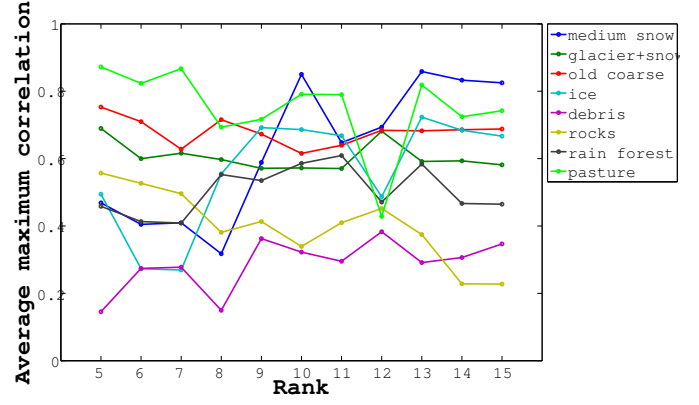


(c)

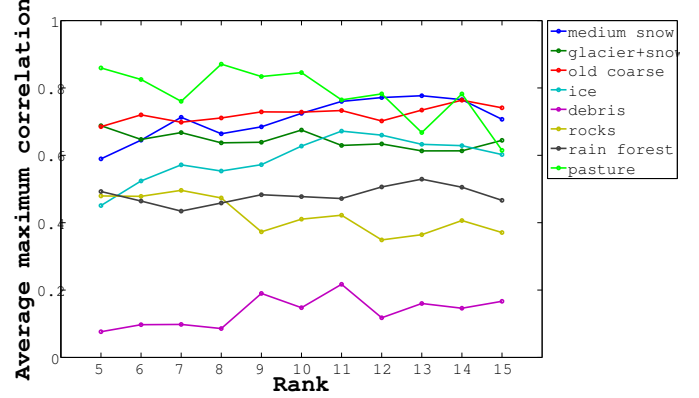
Fig. 5. Average angular distance between the eight library endmembers and the spectral factors obtained by the three competing CP decomposition algorithms: (a) ANLS, (b) CCG and (c) ProCo ALS.

rank values close to $R = 8$, which is an additional evidence that this is a proper rank value for the data tensor. These results can be explained by the high presence of snow and ice throughout all the time series, while it is difficult to visually assess the presence of rocks and debris. Also, the almost flat spectra of these two materials could make it hard for the CP decomposition to select one of them as a spectral factor, since this information could be incorporated to the multilinear decomposition as scaling factors.

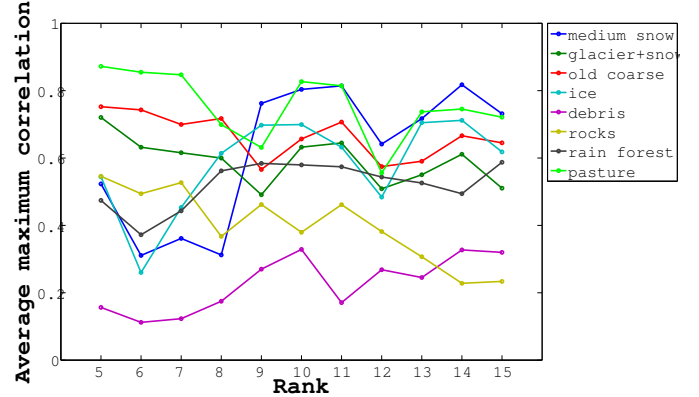
In Fig. 6 we show the average maximum Pearson correlation between the spatial factors estimated by the three CP decomposition algorithms and the spatial abundances obtained



(a)



(b)



(c)

Fig. 6. Average Maximum Pearson correlation between the abundances estimated by FCLSU corresponding to the eight actual endmembers and the spatial factors obtained by the three competing CP decomposition algorithms: (a) ANLS, (b) CCG and (c) ProCo ALS.

by the conventional FCLSU approach. Using the FCLSU, we obtained a set of 8 abundance maps for each of the 44 time acquisitions. We computed the Pearson correlation between the abundance maps and all the spatial factors, and for each run and library endmember, we selected the maximum Pearson correlation value. Then, we computed the mean among the 50 runs for each algorithm and rank value. Again, we obtained high correlation values for the four snow/ice materials, middle correlation values for the *pasture/rain forest*, and small correlations for the *debris/rocks*. Despite the fact that some of

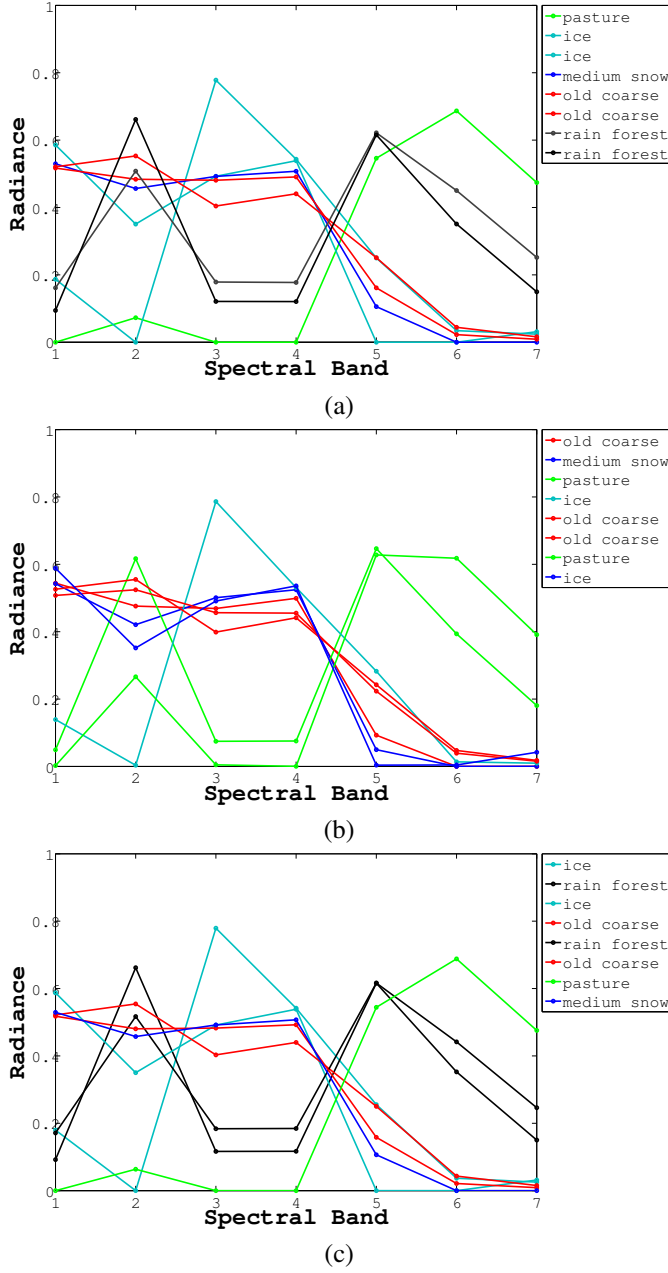


Fig. 7. Spectral factors obtained by the three competing CP decomposition algorithms: (a) ANLS, (b) CCG and (c) ProCo ALS. The colours of the spectral factors correspond to their matching library endmembers.

the materials were not detected by the CP decomposition, the presence of high correlated spatial factors indicates that they are physically meaningful and that the spatial factors could be interpretable in terms of abundances.

2) *Results for $R = 8$* : next, we give details about the CP factors obtained for the best run of each algorithm, in terms of average normalized RMSE, with $R = 8$. Fig. 7 shows the spectral factors obtained by the three CP algorithms. The colors indicate the library endmembers that are the most similar in angular distance to each of the spectral factors. Fig. 8 shows the spectral angular distances between the spectral factors and the library endmembers. It is interesting to mention

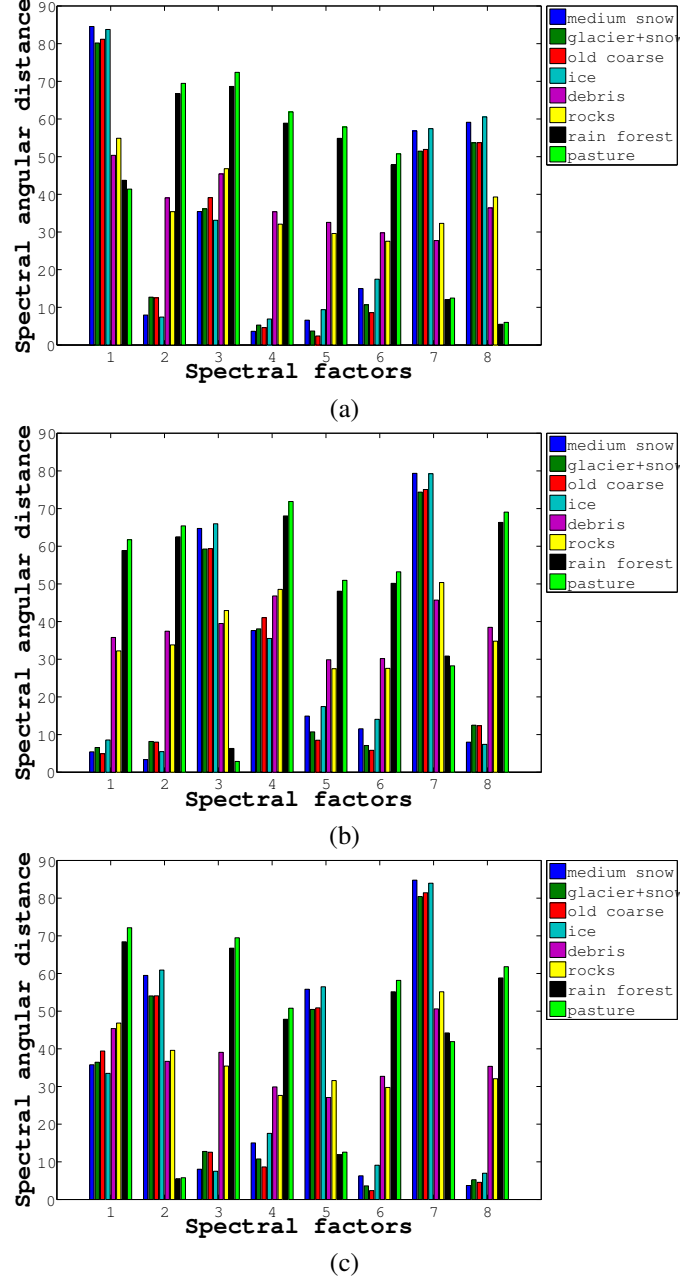


Fig. 8. Spectral angular distances between the library endmembers and the spectral factors obtained by the three competing CP decomposition algorithms: (a) ANLS, (b) CCG and (c) ProCo ALS.

that in all three algorithms, the spectral factors are very similar. The different matchings to the library endmembers are due to slight differences in angular distances, *i.e.* the differences among the four snow/ice spectra are very small. In the three algorithms there are two factors with high minimum spectral angle distances, factors 1 and 3 for the ANLS, factors 4 and 7 for the CCG, and factors 1 and 7 for the ProCo ALS. These factors are being associated with *pasture* and *ice*, and from the shape of the factors (peaks in bands 3 and 6) it can be guessed that they correspond to a mixture of both materials.

Fig. 9 depicts the spatial factors as abundance maps and compare them to the average of the 44 abundance maps

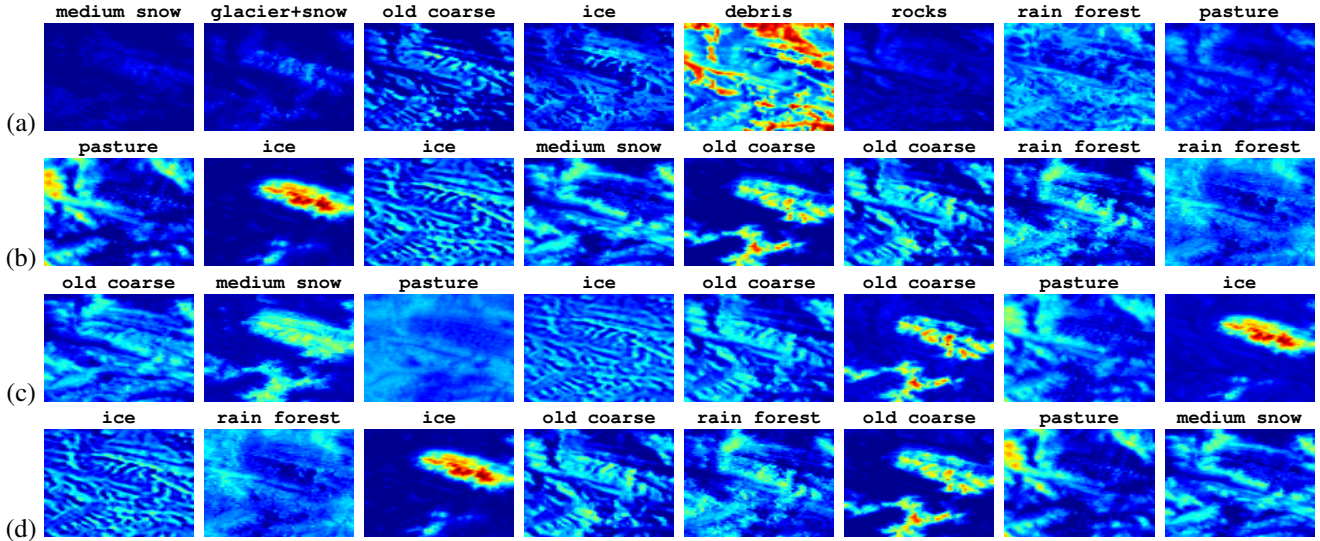


Fig. 9. (a) Spatial abundances obtained by the FCLSU averaged over the 44 time frames, (b-d) Spatial factors obtained by the three competing CP decomposition algorithms: (b) ANLS, (c) CCG and (d) ProCo ALS. Each abundance map is associated to one library endmember and, for each method, scaled by the maximum value of the eight abundance maps.

obtained by the conventional FCLSU approach. The spatial factors obtained by the three CP algorithms are very similar, what evidences that all three algorithms converge to the approximately same local minima given a proper tensor rank. It could be appreciated that two of the factors are related to the permanent snow/ice on the top of the mountain (factors 2 and 5 for ANLS, 6 and 8 for CCG, and 3 and 6 for ProCo ALS). Also, the *rain forest* and *pasture* areas correspond to the low sides of the mountains and surrounding areas. Comparing them to those obtained by the FCLSU, it could be noted that the *debris* FCLSU abundance map has very high values compared to the remaining FCLSU abundance maps. This suggests again that its flat spectra works like a scaling factor, probably modelling spectral variability, which is a well known issue in spectral unmixing [37]. Overall, the estimated spatial factors are meaningful and the qualitative visual assessment encourages us to further investigate the use of the nonnegative CP decomposition as a multilinear “blind” spectral unmixing technique.

Finally, we provide in Fig. 10 the estimated time factors. Same as before, the colors indicate the assignment of the factors to the most similar endmembers. It is difficult to give an interpretation of the factors, but it is possible to show that they are related to the three phases of the snow season: pre-season, snow and post-season. In the pre-season and post-season the *rain forest/pasture* factors have a higher importance and although the *ice/snow* factors are present over all the season due to the permanent ice/snow areas on the top of the mountains, they have higher values during the snow and post-season phases.

V. CONCLUSION

Big data hyperspectral time series or multiangular acquisitions can be represented as tensors. In this paper, we have proposed the use of compression-based nonnegative CP decomposition algorithms to effectively analyze hyperspectral

big data. We showed that the CP decomposition of hyperspectral tensors could be understood as an extension of the linear mixing model to higher-orders, that is, a blind multilinear spectral unmixing technique. We gave experimental evidence that supports the validity of the proposed approach and the physical interpretability of the tensor factors in terms of spectral signatures, fractional spatial abundances and time/angle variations.

Further work will focus in extending the proposed multilinear spectral unmixing approach in several ways: (i) to incorporate physical constraints, as the abundance sum-to-one constraint; (ii) to address nonlinearities in the data, since multilinearity allows to model a large class of nonlinearities; (iii) to incorporate additional regularization terms, usually employed in conventional spectral unmixing, that have been shown to produce enhanced results, such as sparsity and spatial smoothness; and (iv) to better understand the role of the diagonal tensor, \mathcal{L} , to model spectral variability due to scaling factors. Finally, even if far from the scope of the present work, the CP decomposition could be understood as a lossy compression methodology robust to noise, which is an interesting avenue for further research.

ACKNOWLEDGEMENT

Authors would like to thank Prof. Isabelle Zin from the LTHE laboratory, CNRS, Prof. Marie Dumont from Météo-France and Dr. Mauro Dalla Mura from GIPSA-lab, Grenoble INP, for providing the experimental dataset.

REFERENCES

- [1] A. F. H. Goetz, G. Vane, J. E. Solomon, and B. N. Rock, “Imaging spectrometry for earth remote sensing,” *Science*, vol. 228, no. 4704, pp. 1147–1153, 1985.
- [2] A. Plaza, J. A. Benediktsson, J. W. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gualtieri, M. Marconcini, J. C. Tilton, and G. Trianni, “Recent advances in techniques for hyperspectral image processing,” *Remote Sens. Environ.*, vol. 113, Supplement 1, pp. 110–122, 2009.

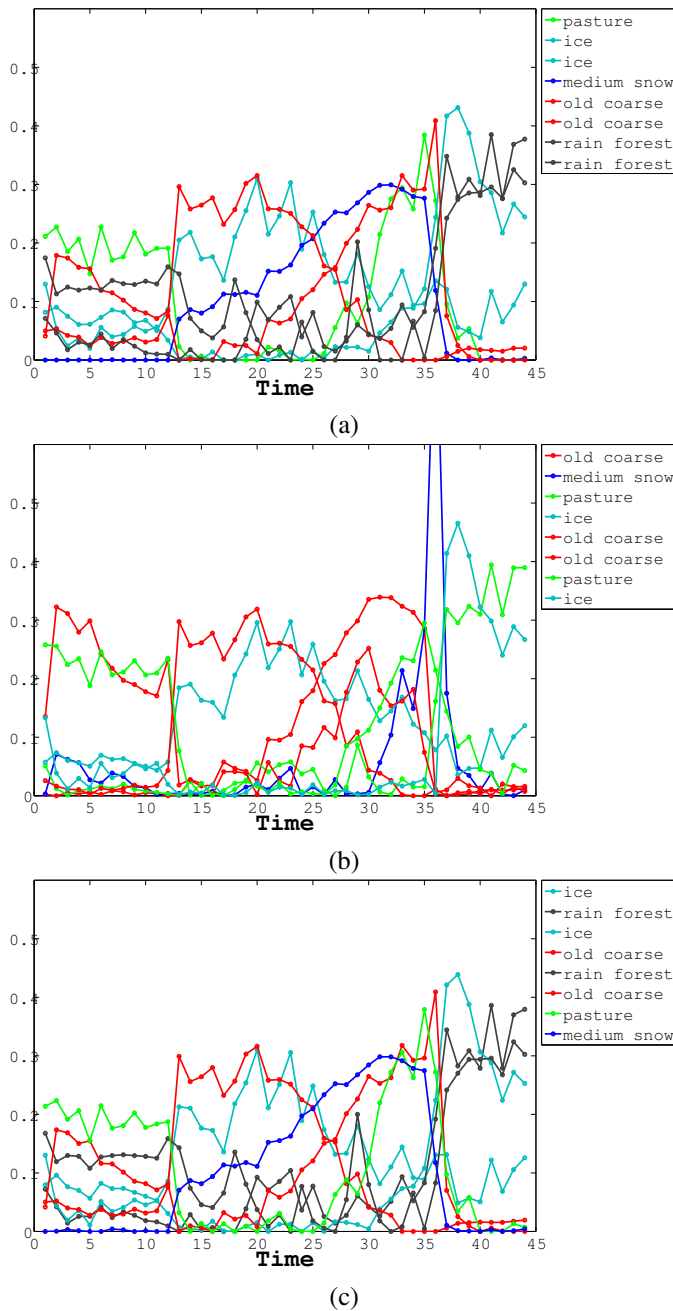


Fig. 10. Time factors obtained by the three competing CP decomposition algorithms: (a) ANLS, (b) CCG and (c) ProCo ALS. The colours of the time factors correspond to their matching library endmembers.

[3] A. F. H. Goetz, "Three decades of hyperspectral remote sensing of the earth: A personal view," *Remote Sens. Environ.*, vol. 113, Supplement 1, pp. 5–16, 2009, imaging Spectroscopy Special Issue.

[4] M. E. Schaepman, S. L. Ustin, A. J. Plaza, T. H. Painter, J. Verrelst, and S. Liang, "Earth system science related imaging spectroscopy an assessment," *Remote Sens. Environ.*, vol. 113, Supplement 1, pp. 123–137, 2009, imaging Spectroscopy Special Issue.

[5] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. M. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, June 2013.

[6] A. Mueller, K. Staenz, C. Ong, and U. Heiden, "International spaceborne imaging spectroscopy (isis) technical committee [technical committees]," *IEEE Geosci. Remote Sens. Mag.*, vol. 2, no. 1, pp. 18–21, March 2014.

[7] K. Frink, L. Hayden, and M. LeCompte, "Compact reconnaissance imaging spectrometer for mars (crism)," in *IEEE Int. Geosci. Remote Sens. Symposium (IGARSS)*, July 2011, pp. 4078–4079.

[8] S. Bourguignon, D. Mary, and E. Slezak, "Restoration of astrophysical spectra with sparsity constraints: Models and algorithms," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 5, pp. 1002–1013, Sept 2011.

[9] A. Cichocki, R. Zdunek, A. H. Phan, and S.-i. Amari, *Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation*. John Wiley & Sons, 2009.

[10] P. Comon, "Tensors : A brief introduction," *IEEE Signal Process. Mag.*, vol. 31, no. 3, pp. 44–53, May 2014.

[11] F. L. Hitchcock, "The expression of a tensor or a polyadic as a sum of products," *J. Math. Physics*, vol. 6, no. 1, pp. 165–189, 1927.

[12] H. A. L. Kiers, "Towards a standardized notation and terminology in multiway analysis," *J. Chemometrics*, pp. 105–122, 2000.

[13] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, D. Qian, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.

[14] D. Roberts, M. Gardner, R. Church, S. Ustin, G. Scheer, and R. Green, "Mapping chaparral in the santa monica mountains using multiple endmember spectral mixture models," *Remote Sens. Environ.*, vol. 65, no. 3, pp. 267 – 279, 1998. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0034425798000376>

[15] N. Dobigeon, S. Moussaoui, M. Coulon, J.-Y. Tourneret, and A. Hero, "Joint bayesian endmember extraction and linear unmixing for hyper-spectral imagery," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4355–4368, Nov 2009.

[16] O. Eches, J. Benediktsson, N. Dobigeon, and J.-Y. Tourneret, "Adaptive markov random fields for joint unmixing and segmentation of hyper-spectral images," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 5–16, Jan 2013.

[17] J. E. Cohen, R. C. Farias, and P. Comon, "Fast decomposition of large nonnegative tensors," *IEEE Signal Process. Lett.*, vol. 22, no. 7, pp. 862–866, 2015.

[18] Q. Zhang, H. Wang, R. J. Plemmons, and V. Pauca, "Tensor methods for hyperspectral data analysis: a space object material identification study," *JOSA A*, vol. 25, no. 12, pp. 3001–3012, 2008.

[19] L. Zhang, L. Zhang, D. Tao, and X. Huang, "Tensor discriminative locality alignment for hyperspectral image spectral-spatial feature extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 242–256, 2013.

[20] D. Letexier, S. Bourennane, and J. Blanc-Talon, "Nonorthogonal tensor matricization for hyperspectral image filtering," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 1, pp. 3–7, 2008.

[21] S. Bourennane, C. Fossati, and A. Cailly, "Improvement of classification for hyperspectral images based on tensor modeling," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 801–805, 2010.

[22] S. Velasco-Forero and J. Angulo, "Classification of hyperspectral images by tensor modeling and additive morphological decomposition," *Pattern Recognit.*, vol. 46, no. 2, pp. 566–577, 2013.

[23] M. Veganzones, J. Cohen, R. Cabral Farias, R. Marrero, C. J., and P. Comon, "Multilinear spectral unmixing of hyperspectral multiangle images," in *European Signal Process. Conf. (EUSIPCO)*, September 2015.

[24] J. Cohen and U. Rothblum, "Nonnegative ranks, decompositions and factorizations of nonnegative matrices," *Lin. Alg. Appl.*, vol. 190, pp. 149–168, 1993.

[25] L.-H. Lim and P. Comon, "Nonnegative approximations of nonnegative tensors," *J. Chemometrics*, vol. 23, no. 7-8, pp. 432–441, 2009.

[26] J. B. Kruskal, "Three-way arrays: Rank and uniqueness of trilinear decompositions," *Lin. Alg. Appl.*, vol. 18, pp. 95–138, 1977.

[27] L. Chiantini, G. Ottaviani, and N. Vannieuwenhoven, "An algorithm for generic and low-rank specific identifiability of complex tensors," *SIAM J. Matrix Ana. Appl.*, vol. 35, no. 4, pp. 1265–1287, 2014.

[28] Y. Qi, P. Comon, and L.-H. Lim, "Uniqueness of nonnegative tensor approximations," Oct. 2014, arxiv:1410.8129.

[29] P. Paatero, "A weighted non-negative least squares algorithm for three-way PARAFAC factor analysis," *Chemometr. Intell. Lab.*, vol. 38, no. 2, pp. 223–242, 1997.

[30] R. Bro, "Multi-way analysis in the food industry: models, algorithms, and applications," Ph.D. dissertation, University of Amsterdam, The Netherlands, 1998.

[31] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM J. Matrix Ana. and Appl.*, vol. 21, no. 4, pp. 1253–1278, 2000.

- [32] N. Sidiropoulos, E. Papalexakis, and C. Faloutsos, "Parallel randomly compressed cubes: A scalable distributed architecture for big tensor decomposition," *IEEE Signal Process. Mag.*, vol. 31, no. 5, pp. 57–70, 2014.
- [33] G. Zhou, A. Cichocki, and S. Xie, "Decomposition of big tensors with low multilinear rank," *arXiv preprint arXiv:1412.1885*, 2014.
- [34] A. P. Liavas and N. D. Sidiropoulos, "Parallel algorithms for constrained tensor factorization via the alternating direction method of multipliers," *arXiv preprint arXiv:1409.2383*, 2014.
- [35] S. J. Wright and J. Nocedal, *Numerical optimization*. Springer New York, 1999, vol. 2.
- [36] J. Bioucas-Dias and M. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, June 2010, pp. 1–4.
- [37] A. Zare and K. Ho, "Endmember variability in hyperspectral analysis: Addressing spectral variability during spectral unmixing," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 95–104, Jan 2014.



Miguel A. Vezanones (M'12) received the M.Sc. and Ph.D. degrees in computer science from the Basque Country University (UPV/EHU), Donostia, Spain, in 2005 and 2012 respectively. Since 2012, he is a postdoctoral researcher at the Images-Signal department in the GIPSA-lab, Grenoble, France. His current main research activities focus in the analysis of hyperspectral images by means of computational intelligence, statistical techniques and tensor data analysis. He is a Reviewer of IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, IEEE JOURNAL OF SELECTED TOPICS IN EARTH OBSERVATIONS AND REMOTE SENSING and the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING.

CTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, IEEE JOURNAL OF SELECTED TOPICS IN EARTH OBSERVATIONS AND REMOTE SENSING and the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING.



Jeremy E. Cohen received a licence in mathematics at the University of Lyon, and graduated from the Ecole Centrale de Lyon with both a mathematical engineering master degree and a research master degree in telecommunications in 2013. He worked at Imai's Laboratory at the University of Tokyo on quantum computation applied to communication. He is currently a Ph.D student at GIPSA-Lab under the supervision of Pierre Comon. His domains of interest are tensor decomposition applied to environmental data mining, but also information theory, optimization and quantum computation.

tion and quantum computation.



Rodrigo Cabral Farias was born in Porto Alegre, Brazil, in 1986. He received the B.Sc. degree in electrical engineering from the Federal University of Rio Grande do Sul (UFRGS), Porto Alegre, Brazil, and from the Grenoble Institute of Technology (Grenoble-INP), Grenoble, France, both in 2009. He received the M.Sc and Ph.D. degree in signal processing from Grenoble-INP and Grenoble University in 2009 and 2013 respectively. He is currently working as a post doctoral researcher at the GIPSA-Lab (Grenoble Laboratory of Image,

Speech, Signal, and Automation). His research concerns statistical signal processing, digital communications, sensor networks and tensor methods for signal processing.



Jocelyn Chanussot (M'04-SM'04-F'12) received the M.Sc. degree in electrical engineering from the Grenoble Institute of Technology (Grenoble INP), Grenoble, France, in 1995, and the Ph.D. degree from Savoie University, Annecy, France, in 1998. In 1999, he was with the Geography Imagery Perception Laboratory for the Delegation Generale de l'Armement (DGA - French National Defense Department). Since 1999, he has been with Grenoble INP, where he was an Assistant Professor from 1999 to 2005, an Associate Professor from 2005

to 2007, and is currently a Professor of signal and image processing. He is conducting his research at the Grenoble Images Speech Signals and Automatics Laboratory (GIPSA-Lab). His research interests include image analysis, multicomponent image processing, nonlinear filtering, and data fusion in remote sensing. He is a member of the Institut Universitaire de France (2012-2017). Since 2013, he is an Adjunct Professor of the University of Iceland. Dr. Chanussot is the founding President of IEEE Geoscience and Remote Sensing French chapter (2007-2010) which received the 2010 IEEE GRS-S Chapter Excellence Award. He was the co-recipient of the NORSIG 2006 Best Student Paper Award, the IEEE GRSS 2011 Symposium Best Paper Award, the IEEE GRSS 2012 Transactions Prize Paper Award and the IEEE GRSS 2013 Highest Impact Paper Award. He was a member of the IEEE Geoscience and Remote Sensing Society AdCom (2009-2010), in charge of membership development. He was the General Chair of the first IEEE GRSS Workshop on Hyperspectral Image and Signal Processing, Evolution in Remote sensing (WHISPERS). He was the Chair (2009-2011) and Cochair of the GRS Data Fusion Technical Committee (2005-2008). He was a member of the Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Society (2006-2008) and the Program Chair of the IEEE International Workshop on Machine Learning for Signal Processing, (2009). He was an Associate Editor for the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS (2005-2007) and for PATTERN RECOGNITION (2006-2008). Since 2007, he is an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. Since 2011, he is the Editor-in-Chief of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING. In 2013, he was a Guest Editor for the PROCEEDINGS OF THE IEEE and in 2014 a Guest Editor for the IEEE SIGNAL PROCESSING MAGAZINE. He is a Fellow of the IEEE and a member of the Institut Universitaire de France (2012-2017).



Pierre Comon (M'87 – SM'95 – F'07) graduated in 1982, and received the Doctorate degree in 1985, both from the University of Grenoble, France. He later received the Habilitation to Lead Researches in 1995, from the University of Nice, France. He has been for nearly 13 years in industry, first with Crouzet-Sextant, Valence, France, between 1982 and 1985, and then with Thomson Marconi, Sophia-Antipolis, France, between 1988 and 1997. He spent 1987 with the ISL laboratory, Stanford University, CA. He joined in 1997 the Eurecom

Institute, Sophia-Antipolis, France. He is research director with CNRS since 1998, first at laboratory I3S, Sophia-Antipolis, France, until 2012, and then at Gipsa-Lab, Grenoble, France. His research interests include Blind techniques, Statistical Signal and Array Processing, Tensor decompositions, Multi-Way Factor Analysis and its applications to biomedical end environment.

Dr. Comon was Associate Editor of the *IEEE Transactions on Signal Processing* from 1995 to 1998, and a member of the French National Committee of Scientific Research from 1995 to 2000. He was the coordinator of the European Basic Research Working Group on High-Order Statistics, *ATHOS*, from 1992 to 1995. Between 1992 and 1998, he was a member of the Technical and Scientific Council of the Thomson Group. Between 2001 and 2004 he acted as launching Associate Editor with the *IEEE Transactions on Circuits and Systems I*, in the area of Blind Techniques. He has also been a member of the editorial board of the Elsevier journal *Signal Processing*, and member of several IEEE TC. He is presently in the editorial board of *SIAM Journal on Matrix Analysis and Applications*.